Bi-criterion Shortest Path Problem with a General Non-additive Cost

Peng (Will) Chen and Yu (Marco) Nie

Department of Civil and Environmental Engineeering
Northwestern University

July 19, 2013

NORTHWESTERN UNIVERSITY

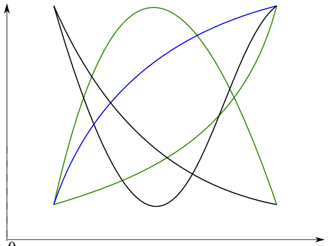
Introduction

A variant of shortest path problem

The problem seeks to optimize a combination of two path attributes, one of which is evaluated by a nonlinear function, i.e.

$$minimize\ P_1^k + h(P_2^k)$$

where P_i^k $(i = 1, 2)$ is ith property of path k and h is a general nonlinear function.



Nonlinear function of one path attribute

Introduction

Literature review

- Dial (1979) proposes an algorithm which can solve the shortest path problem with the linear combination of two path attributes.
- Henig (1985) uses a line search method to find the path that admits the best upper bound and further close the gap with a K-shortest path search.
- Mirchandani & Wiecek (1993) refines Henig's linear search method.
- Tsaggouris & Zaroliagis (2004) provides a two-phase algorithm.
- ......

Introduction

Literature review

- Dial (1979) proposes an algorithm which can solve the shortest path problem with the linear combination of two path attributes.
- Henig (1985) uses a line search method to find the path that admits the best upper bound and further close the gap with a K-shortest path search.
- Mirchandani & Wiecek (1993) refines Henig's linear search method.
- Tsaggouris & Zaroliagis (2004) provides a two-phase algorithm.
- ......

Main differences from existing work

- The general nonlinear function;
- Efficient partial path enumeration;
- Graphical illustration.

Applications

Nonlinear traffic assignment problem

- Nonlinear valuation of travel time and emissions (Gabriel & Bernstein 1997)

Applications

Nonlinear traffic assignment problem

- Nonlinear valuation of travel time and emissions (Gabriel & Bernstein 1997)

Distance-based congestion pricing

- Nonlinear toll function (Lawphongpanich & Yin 2012)

Applications

Nonlinear traffic assignment problem

- Nonlinear valuation of travel time and emissions (Gabriel & Bernstein 1997)

Distance-based congestion pricing

- Nonlinear toll function (Lawphongpanich & Yin 2012)

Tradable mobility credit problem

- Nonlinear toll function (Yang & Wang 2011, Nie 2012)

Applications

Nonlinear traffic assignment problem

- Nonlinear valuation of travel time and emissions (Gabriel & Bernstein 1997)

Distance-based congestion pricing

- Nonlinear toll function (Lawphongpanich & Yin 2012)

Tradable mobility credit problem

- Nonlinear toll function (Yang & Wang 2011, Nie 2012)

Optimal path problem considering schedule penalty

- Nonlinear schedule cost (Nie et al. 2011)

Applications

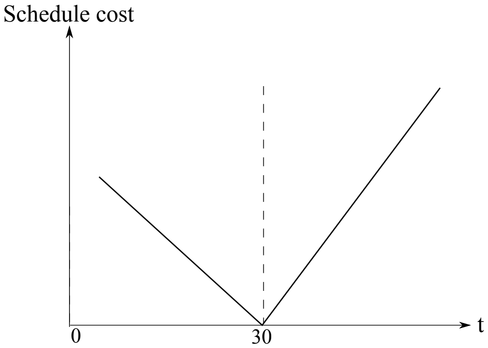
A SIMPLE EXAMPLE

Assume that a traveler departs from home at 8:00 AM. His desired arrival time at the workplace is 8:30 AM.

Applications

A SIMPLE EXAMPLE

Assume that a traveler departs from home at 8:00 AM. His desired arrival time at the workplace is 8:30 AM.



Schedule cost:a non-monotone function of travel time

Problem Formulation

Formulation

The problem is finding the path between an $O - D$ pair $r - s$ to

$$\text{minimize } P_1^k + h(P_2^k), \text{ subject to: } k \in K \tag{1}$$

where P_i^k $(i = 1, 2)$ is ith property of path k and h is a general nonlinear function.

Problem Formulation

Formulation

The problem is finding the path between an $O - D$ pair $r - s$ to

$$\text{minimize } P_1^k + h(P_2^k), \text{ subject to: } k \in K \tag{1}$$

where P_i^k $(i = 1, 2)$ is ith property of path k and h is a general nonlinear function.

Two possible interpretations

- P_1^k as monetary cost c_k and P_2^k as travel time t_k, the function h can be considered as an evaluation of travel time in the monetary cost;
- P_1^k as travel time t_k and P_2^k as travel distance l_k, the function h can be considered as a distance-based toll measured in the unit of time.

Problem Formulation

Formulation

The problem is finding the path between an $O-D$ pair $r-s$ to

$$\text{minimize } P_1^k + h(P_2^k), \text{ subject to: } k \in K \tag{1}$$

where P_i^k $(i = 1, 2)$ is ith property of path k and h is a general nonlinear function.

Two possible interpretations

- P_1^k as monetary cost c_k and P_2^k as travel time t_k, the function h can be considered as an evaluation of travel time in the monetary cost;
- P_1^k as travel time t_k and P_2^k as travel distance l_k, the function h can be considered as a distance-based toll measured in the unit of time.

Note

Hereafter, we shall consider h as a function of t_k and c_k is the other path cost.

Algorithm Outline

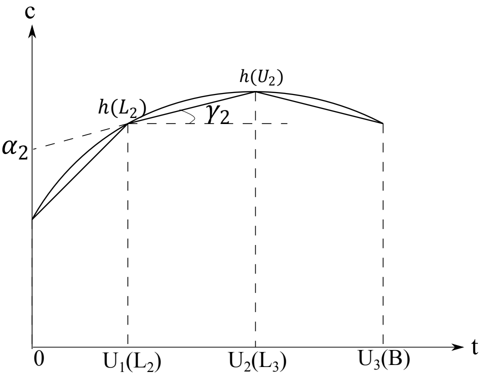
<u>Basic idea</u>

1. Approximate the nonlinear function $h(\cdot)$ with a piecewise linear function;
2. Decompose the problem into several sub-problems associated with each linear piece;
3. Solve each sub-problem sequentially;
4. Report the optimal solution or the best upper bound to the linearized problem.

Approximating nonlinear function $h(t)$

A piecewise linear function $H(t)$

First, the feasible range for t is divided into m intervals as $[L_j, U_j]$ where $L_1 = 0$, $U_m = B$, and $L_j = U_{j-1}$ for $j = 2, ..., m$.



Approximating nonlinear function $h(t)$

A piecewise linear function $H(t)$

For each segment j, the slope of the line, denoted as γ_j, can be obtained as

$$\gamma_j = \frac{h(U_j) - h(L_j)}{U_j - L_j} \tag{2}$$

The line intersects with the vertical axis $(t = 0)$ at

$$\alpha_j = h(L_j) - \gamma_j L_j \tag{3}$$

Therefore, we can write the piecewise linear function $H(t)$ as follows:

$$H(t) = \alpha_j + \gamma_j t; \quad \text{if } t \in [L_j, U_j], j = 1, ..., m \tag{4}$$

Decomposition

Due to discretization, the linearized problem can be decomposed into a sequence of subproblems as follows:

$$\min_{k \in K} z_j = c_k + \gamma_j t_k \tag{5a}$$

$$\text{subject to: } t_k \in [L_j, U_j] \tag{5b}$$

Then, the optimal solution to linearized Problem (1) can be found by solving the following problem:

$$z = \min_{j=1,\ldots,m} \alpha_j + z_j^* \tag{6}$$

where z_j^* is the optimal solution to the jth subproblem.

Decomposition

Due to discretization, the linearized problem can be decomposed into a sequence of subproblems as follows:

$$\min_{k \in K} z_j = c_k + \gamma_j t_k \quad (5a)$$

$$\text{subject to: } t_k \in [L_j, U_j] \quad (5b)$$

Then, the optimal solution to linearized Problem (1) can be found by solving the following problem:

$$z = \min_{j=1,\dots,m} \alpha_j + z_j^* \quad (6)$$

where z_j^* is the optimal solution to the jth subproblem.

Note

The subproblem (5) is a constrained shortest path problem with a linear objective function $g_k = c_k + \gamma t_k$.

Efficient Path Set

Suppose now that travellers would choose paths based on a general cost with a linear function h, defined by

$$g_k = c_k + h(t_k) = c_k + \gamma t_k, \tag{7}$$

where $\gamma \in \mathbb{R}$ is a real scalar that converts travel time to an equivalent monetary cost. An efficient path is formally defined as follows in this paper.

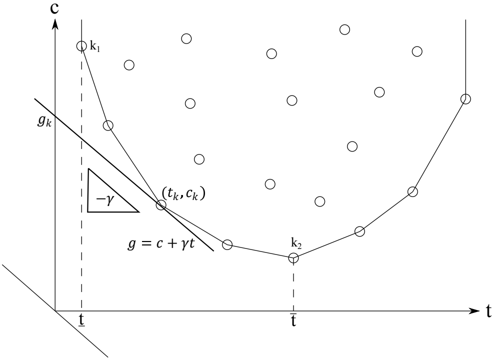
Definition (Efficient Path)

A path k is efficient if (1) it is simple, i.e. it does not contain any cycles and (2) for some $\gamma \in \mathbb{R}$, there exists no other simple path k' such that $g_{k'} < g_k$.

Simply speaking, an efficient path k must have minimum cost g_k for some γ among all simple paths between the O-D pair.

Efficient Path Set

Graphical illustration of efficient paths



Notations

- E_{rs}: the set of efficient paths
- $K^1 = \text{argmin}\{t_k, k \in K\}$
- $K^2 = \text{argmin}\{c_k, k \in K\}$
- $k_1 = \text{argmin}\{c_k, k \in K^1\}$
- $k_2 = \text{argmin}\{t_k, k \in K^2\}$
- $\underline{t} \equiv t_{k_1}, \bar{t} \equiv t_{k_2}$
- $E^+_{rs} = \{k | t_k \leq \bar{t}, k \in E_{rs}\}$
- $E^-_{rs} = \{k | t_k > \bar{t}, k \in E_{rs}\}$

Note

- Given E_{rs}, γ, a minimum cost simple path can be easily identified
- Dial-Henig algorithm for generating E^+_{rs}

Solve the subproblem

Description & Justification

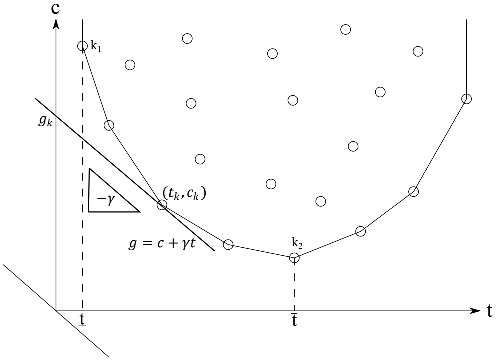
According to the feasible time interval of the subproblem $[L_j, U_j]$ and two special time points $\underline{t}$ and $\overline{t}$, the subproblem could be divided into four cases:

Case 0: $L_j < U_j < \underline{t}$

Case 1: $\underline{t} \leq U_j \leq \overline{t}$

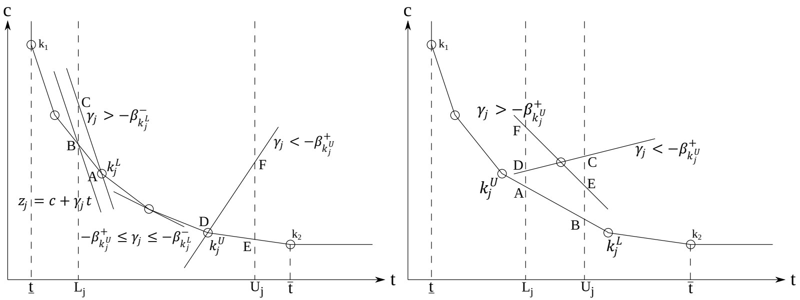
Case 2: $L_j \leq \overline{t} < U_j$

Case 3: $\overline{t} < L_j < U_j$



Solve the subproblem

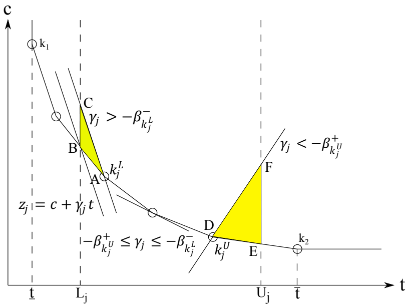
ILLUSTRATION OF CASE 1



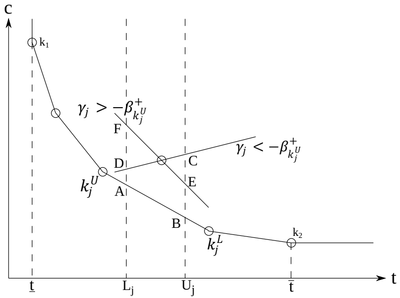
(a) Efficient path exists within interval j (b) Efficient path does not exist within interval j

Solve the subproblem

ILLUSTRATION OF CASE 1



(a) Efficient path exists within interval j



(b) Efficient path does not exist within interval j

NORTHWESTERN UNIVERSITY

Solve the master problem

- After solving each sub-problem, we update the lower bound and the best upper bound of the master problem;
- In some cases, the optimal solution to the sub-problem cannot be found without path enumeration. When this happens, we will first check if the current subproblem has a chance to improve the solution of the master problem, then decide if the path enumeration is necessary.
- Once all sub-problems are solved, we can report the best upper bound of the master problem and the gap.

Analytical results

Theorem 1

Let k^* be an optimal solution to Problem (1) linearized with a piecewise linear function $H(\cdot)$. If $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_m$, i.e. $H(\cdot)$ is concave, then $k^* \in E_{rs}$.

Corollary 1

Let k^* be an optimal solution to Problem (1) linearized with a piecewise linear function $H(\cdot)$. If $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_m \geq 0$, i.e. $H(\cdot)$ is concave and non-decreasing, then $k^* \in E_{rs}^+$.

The proof is referred to the paper.

Note that the corollary verifies the result given in Henig (1985) and Mirchandani & Wiecek (1993), which consider concave and monotone functions.

Overview

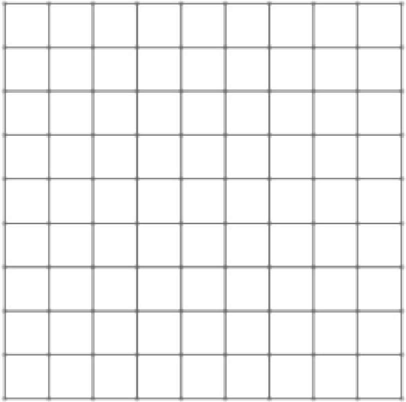
- Coded using TNM, a C++ library for network applications (Nie 2006)
- Tested on a laptop with Window 7 Home Premium, Intel(R) Core(TM) i7-2630QM CPU@2.00GHz and 8.00 GB memory
- Three classes of problems are tested:
  - a small textbook example
  - ten by ten grid networks
  - a large real-world transportation network — Chicago Regional Network

  For the two latter classes of networks, link properties are randomly generated using Gamma distribution. The mean, variance and minimum value of both travel time and cost are 2.5, 5 and 0.5, respectively.

Grid Networks

Overview

The chosen $O - D$ pair for the grid network is from the left-bottom node to the right-top node. Both acyclic and cyclic networks are tested. The acyclic network is included in the test mainly because it allows us to compare the best solution given by the algorithm with the true optimal solution obtained from the brute-force path enumeration.

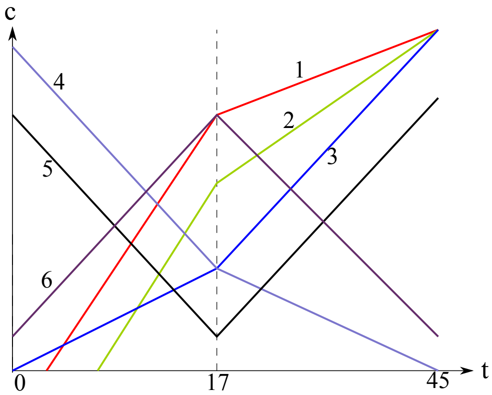


Topology of the 10×10 grid network

Grid Networks

Different types of piece-wise linear functions

The piece-wise linear functions we tested include increasing functions, decreasing functions and non-monotone functions ("V" shape and "Λ" shape).



Grid Networks

Numerical Results of Acyclic Grid Networks

Numerical results of two-piece linear functions for the 10 $\times$ 10 acyclic grid network

	Pieces($[L_j, U_j]$)	Slopes(γ_j)	Best Obj.	Gap	Enum. Paths	Y	Optimal Obj.
1	[0,17];[17,45]	4;1	74.6680	0	0	1000	74.6680
2	[0,17];[17,45]	4;3	74.6680	0	0	1000	74.6680
3	[0,17];[17,45]	3;4	60.7197	0	0	1000	60.7197
4(a)	[0,17];[17,45]	-4;-3	-115.7760	23.4594	4000	2000	-121.7880
4(b)	[0,17];[17,45]	-4;-3	-118.6350	20.5998	6000	3000	-121.7880
4(c)	[0,17];[17,45]	-4;-3	-121.7880	17.4469	11604	7000	-121.7880
4(d)	[0,17];[17,45]	-4;-3	-121.7880	0	53224	50000	-121.7880
5(a)	[0,17];[17,45]	-4;4	-52.3360	0.9586	1295	1000	-52.3360
5(b)	[0,17];[17,45]	-4;4	-52.3360	0	4899	5000	-52.3360
6(a)	[0,17];[17,45]	4;-4	-1.3766	29.8584	2000	1000	-10.1825
6(b)	[0,17];[17,45]	4;-4	-6.2832	24.9519	4000	2000	-10.1825
6(c)	[0,17];[17,45]	4;-4	-10.1825	21.0526	8000	4000	-10.1825
6(d)	[0,17];[17,45]	4;-4	-10.1825	0	48620	50000	-10.1825

Grid Networks

Numerical Results of Acyclic Grid Networks

Approximate the nonlinear function of the following forms:

(1) $a(x-b)^2+c$;

(2) $ae^{bx}+c$.

Numerical results of nonlinear functions for the 10×10 acyclic grid network

	Function	Pieces($[L_j, U_j]$)	Approx. Obj.	Enum. Paths	Y	Optimal Obj.	Gap
1	x^2	[0,20],[20,45]	206.8788	0	50000	206.8788	0
2	$0.1x^2$	[0,20],[20,45]	38.3303	0	50000	38.3303	0
3	$10x^2$	[0,20],[20,45]	1884.9188	0	50000	1884.9188	0
4	$(x-20)^2$	[0,20],[20,45]	15.0206	97240	50000	13.2312	1.7894
5	$-(x-20)^2$	[0,20],[20,45]	-451.0271	48620	50000	-451.8664	0.8393
6	$e^{0.1x}$	[0,20],[20,45]	19.3682	0	50000	19.3682	0
7	$e^{0.01x}$	[0,20],[20,45]	14.0255	2	50000	14.0255	0
8	$-e^{0.1x}$	[0,20],[20,45]	-43.3725	48632	50000	-43.3725	0
9	$-e^{0.01x}$	[0,20],[20,45]	11.4818	8	50000	11.4818	0
10	$e^{-0.1x}$	[0,20],[20,45]	12.8476	4	50000	12.8476	0
11	$e^{-0.01x}$	[0,20],[20,45]	13.5443	4	50000	13.5443	0

Large Scale Real Transportation Network

OVERVIEW

A large scale real transportation network, the Chicago Regional network (Bar-Gera et al. 2010), is used to test the computational performance of the proposed algorithm. The network has 12,982 nodes and 39,018 links. All the link cost and time are randomly generated.

Topology of the Chicago Regional network

Large Scale Real Transportation Network

Numerical Results

Numerical results of two-piece linear functions for the Chicago Regional network

	Pieces($[L_j, U_j]$)	Slopes(γ_j)	Best Obj.	Gap	Enum. Simple Paths	Y^*	CPU Time (s)
1	[0,60];[60,120]	4;1	273.5131	0	0	1000	0.0130
2	[0,60];[60,120]	4;3	273.5131	0	0	1000	0.0150
3	[0,60];[60,120]	3;4	219.6575	0	0	1000	0.0160
4	[0,60];[60,120]	0;3	53.3869	0.0665	879	1000	32.3070
5	[0,60];[60,120]	-4;-3	-259.6078	107.6952	1254	1000	49.0000
6	[0,60];[60,120]	-4;-1	-210.5143	36.7887	1254	1000	48.3290
7	[0,60];[60,120]	-4;4	-185.7762	0.9034	1341	1000	34.0860
8	[0,60];[60,120]	4;-4	195.8455	143.1485	790	1000	33.1340

Conclusions

Main findings from numerical experiments:

- The performance of the algorithm is satisfactory for increasing and "V" shape functions.
- For the decreasing and "Λ" shape functions, extensive path enumeration is often needed.
- Piecewise linear functions with two or three segments seem to provide good approximation to the nonlinear cost functions tested in our experiments (quadratic and exponential).

Questions or comments?

pengchen@u.northwestern.edu

www.pw-chen.com